MetaboAnalyst 5.0

A Web-based Tool for Streamlined Metabolomics Data Analysis

2022.07.12

4. Joint Pathway Analysis

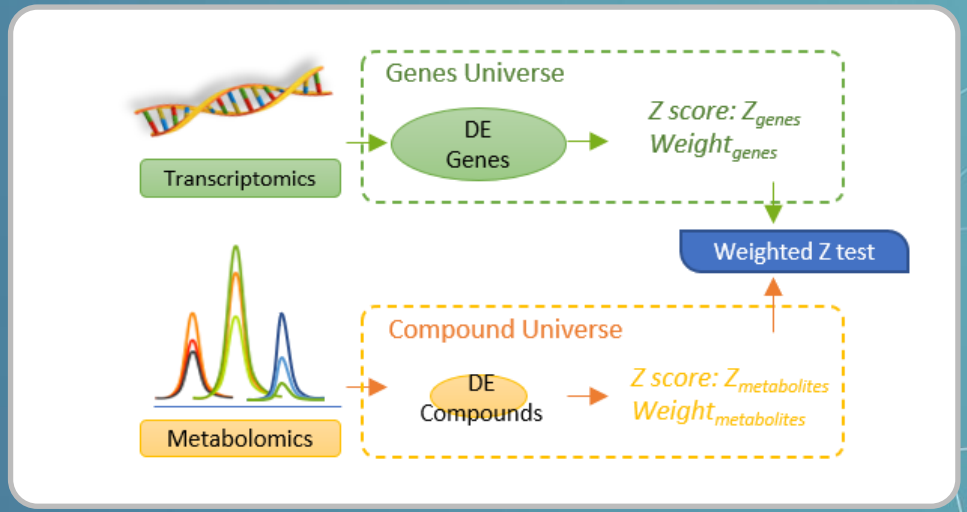
The Joint Pathway Analysis module of MetaboAnalyst has added multiple enhancements for Version 5.

Highlights:

- Added support for an additional 22 model organisms.
- A weighted Z-tests based joint pathway analysis algorithm has been developed to address the issue of genes overwhelming integrated pathway analysis results.
- Multiple integration strategies have been implemented to further enhance the functional interpretation of omics data.
- Support integration for both targeted and untargeted metabolomics.

4.0 Knowledge & Background

- Joint Pathway Analysis is based on weighted integration. The weight strategy for different 'universes' ("transcriptomics data universe" and "metabolomics data universe") are performed with a weighted z-test. The weighted z-test is proposed for combination analysis by Dmitri V. Zaykin and leveraged here for the weighted integration of different datasets with significantly different sizes. Here we specifically prepared a figure to clearly illustrate the mechanism of this weighted integration of different Omics-data in MetaboAnalyst below.



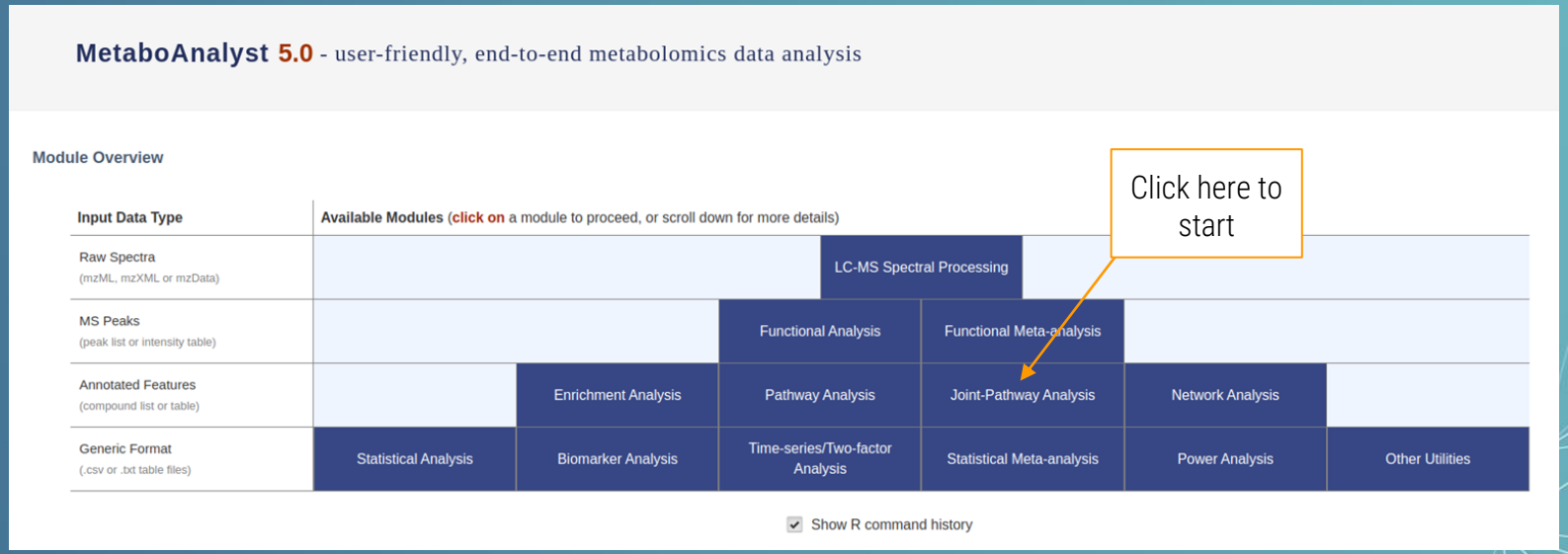
4.0 Knowledge & Background

Specifically, different weights are assigned based on the proportion of genes and metabolites in the specific ‘omics ‘universes’ to balance the influence from the different sizes of the ‘omics inputs upon the integrated pathway results. The adjusted *P* value is estimated with a weighted Z-test as below,

$$P\_merged = 1 - \Phi\left(\frac{\sum_{i=1}^{2} w_i Z_i}{\sqrt{\sum_{i=1}^{2} w_i^2}}\right)$$

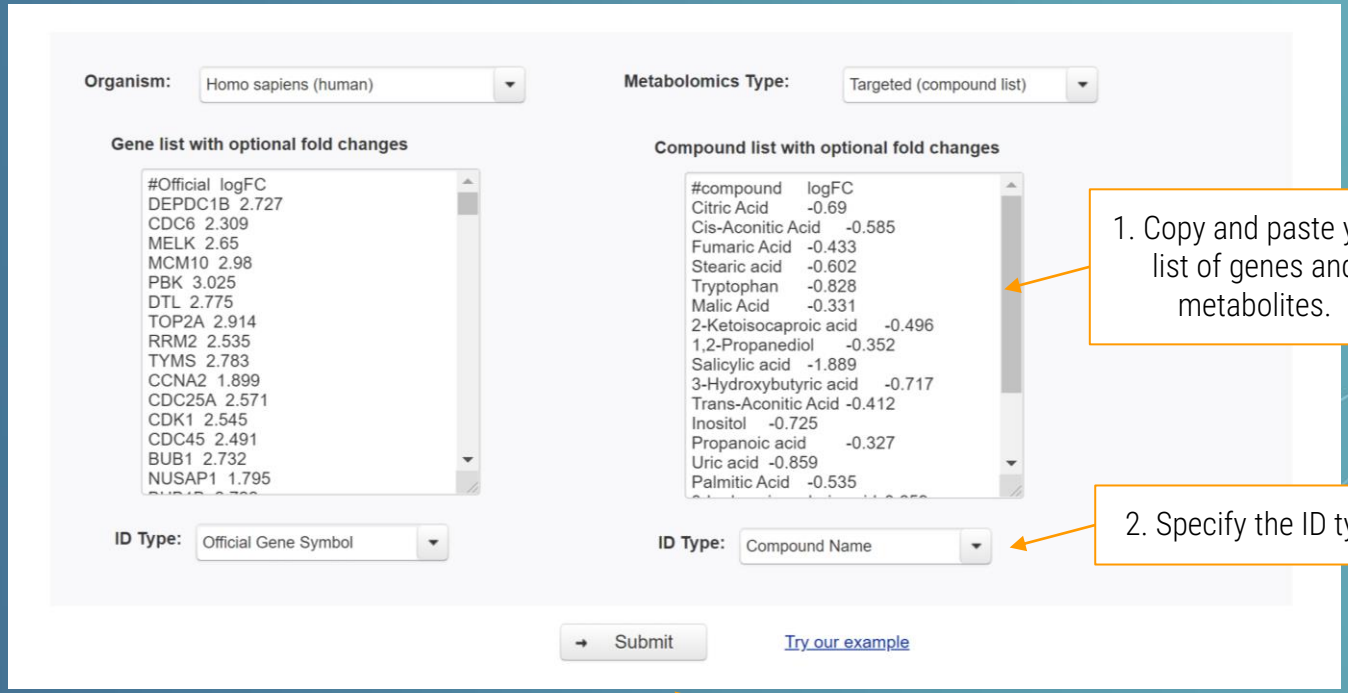
Where w_i is the weights of the *P* values of genes or compounds within individual omics "universe" or "pathway space", respectively; Z_i is the Z score of the corresponding *P* values of single omics data, usually, $Z_i = \Phi^{-1}(1 - P_i)$; P_i is the *P* values from the enrichment analysis above; Φ denote the standard normal cumulative distribution function.

4.1 Start Joint Pathway Analysis



4.2 Integration of genes/proteins with compounds

4.2.1 Data Upload (genes and compounds)



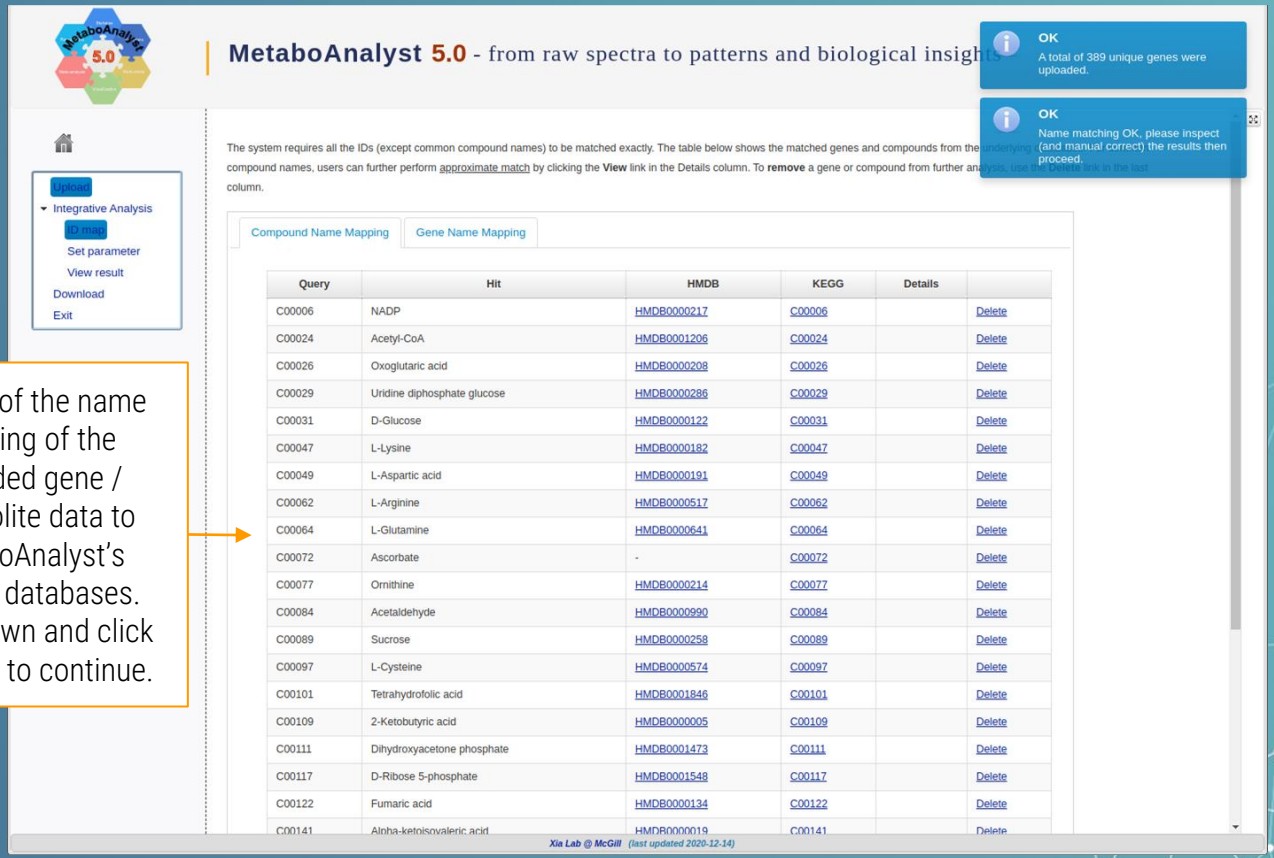
TIP: The Fold change is optional. The titles of the 2 columns need to start with '#'.

1. Copy and paste your list of genes and metabolites.

2. Specify the ID type.

3. Click "Submit" to upload the data.

4.2.2 Name Matching Results

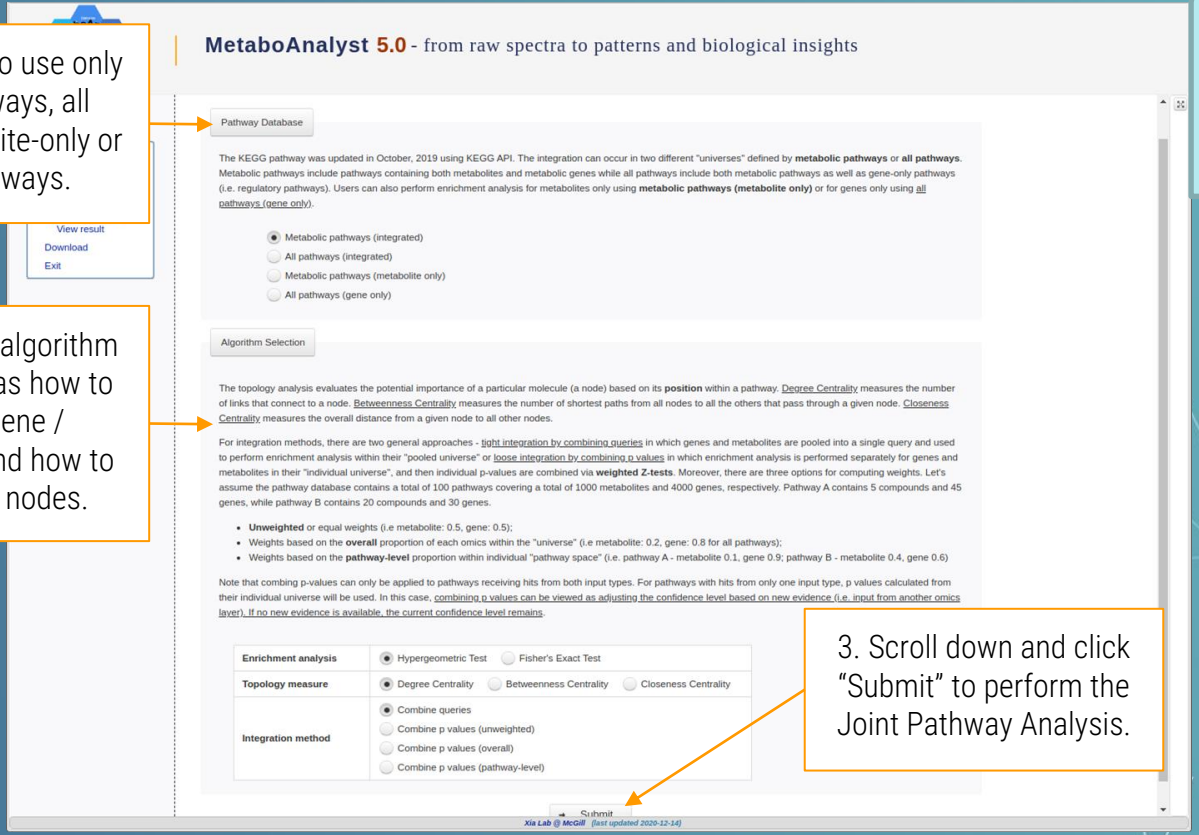


Results of the name mapping of the uploaded gene / metabolite data to MetaboAnalyst's internal databases. Scroll down and click "**Submit**" to continue.

4.2.3 Parameter Selection

1. Select whether to use only metabolic pathways, all pathways, metabolite-only or gene-only pathways.

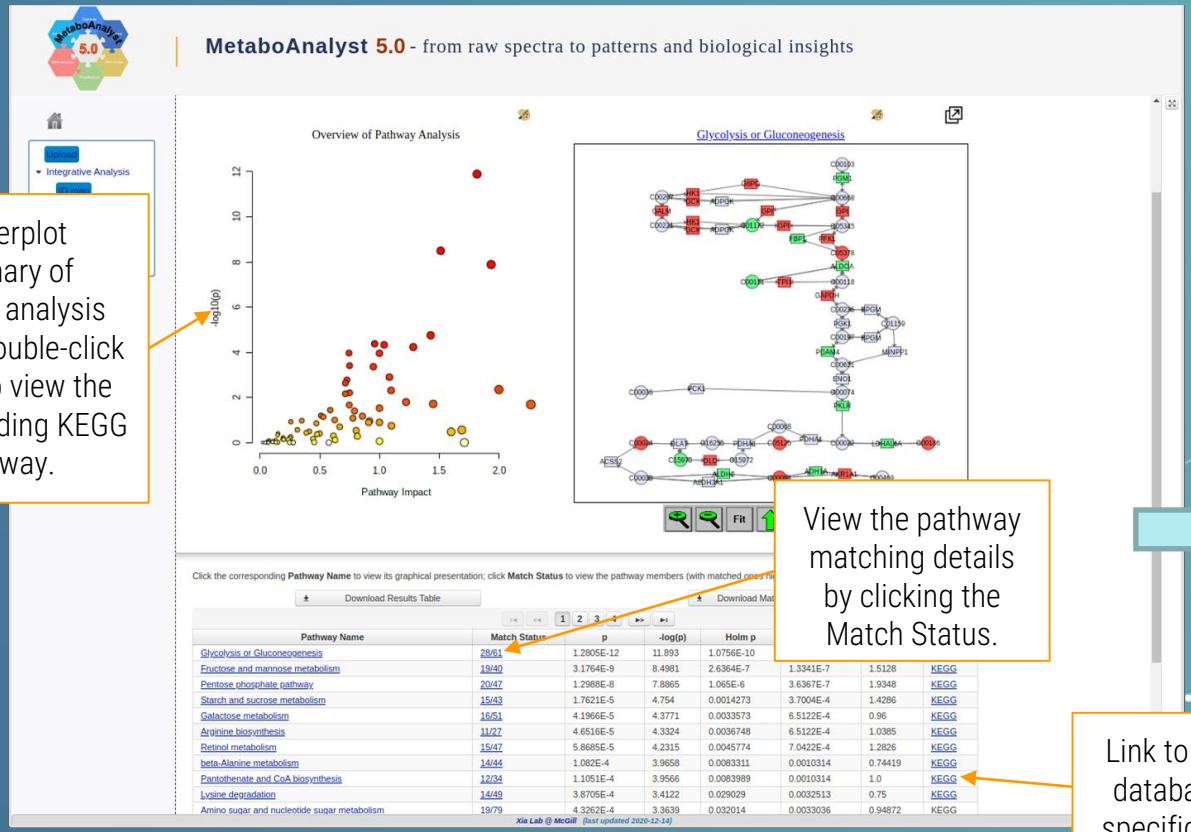
2. Next, decide on algorithm parameters such as how to combine the gene / metabolite data and how to score important nodes.



3. Scroll down and click "Submit" to perform the Joint Pathway Analysis.

TIP: Please carefully read the instructions of the different options of algorithms before submitting your data.

4.2.4 Integration Results



Scatterplot summary of pathway analysis results. Double-click a node to view the corresponding KEGG pathway.

View the pathway matching details by clicking the Match Status.

Link to the KEGG database of the specific pathway.

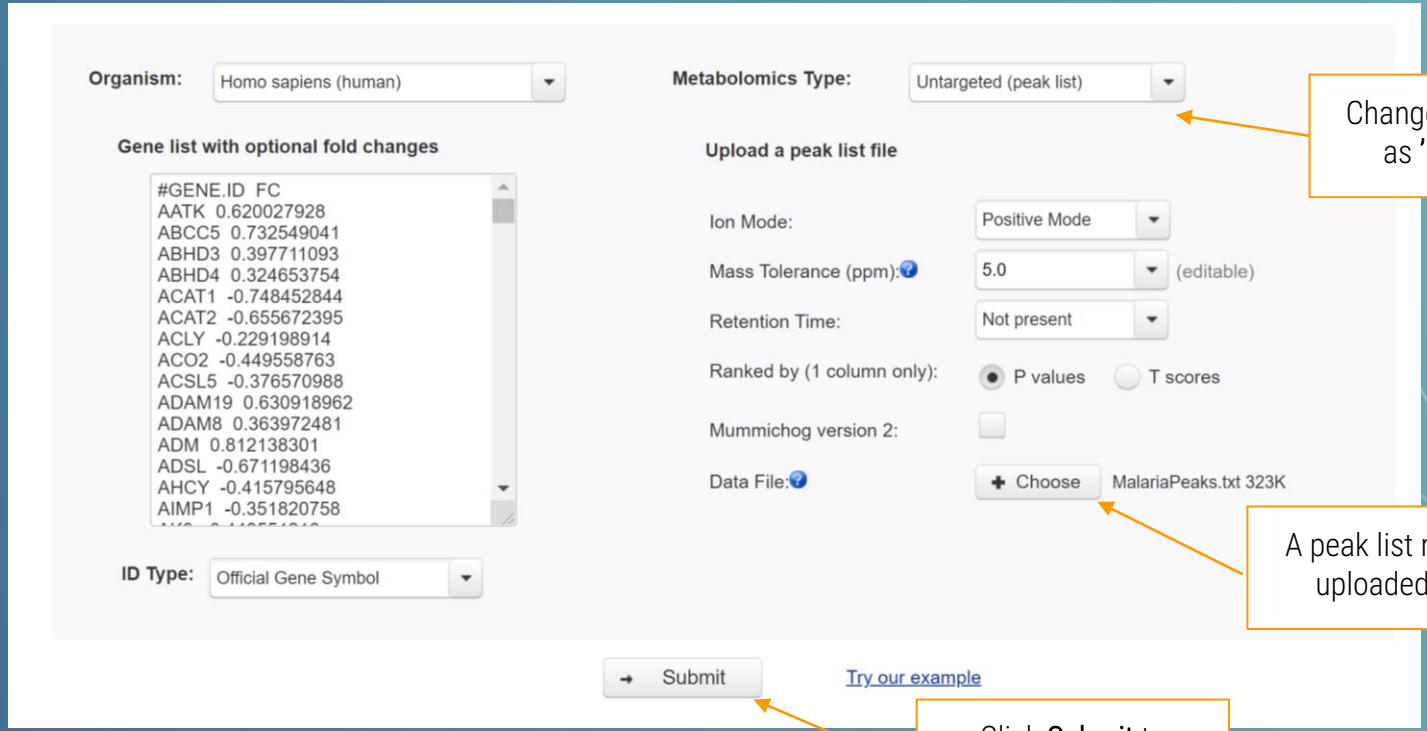
Matched Features

Pathway	Members
Fructose and mannose metabolism	D-Sorbitol; D-Fructose; D-Mannose; beta-D-Fructose 2,6-bisphosphate; D-Mannose 6-phosphate; L-Fucose 1-phosphate; GDP-4-dehydro-6-deoxy-D-mannose; GDP-mannose; D-Mannose 1-phosphate; beta-D-Fructose 6-phosphate; D-Glyceraldehyde 3-phosphate; D-Fructose 1-phosphate; D-Glyceraldehyde; beta-D-Fructose 1,6-bisphosphate; 6-Deoxy-L-galactose; L-Fuconate; alpha-D-Glucose; GDP-L-fucose; Glycerone phosphate; 2-Dehydro-3-deoxy-L-fuconate; AKR1B1, ADR, ALDR1, ALR2, AR...; SORD, HEL-S-95n, RDH, SDH, SORD1, XDH; KHK; HK1, HK, HK1-ta, HK1-tb, HK1-tc, HKD, HKI, HMSNR, HXK1, RP79, hexokinase...; PFKFB1, F6PK, HL2K, PFRX...; MPI, CDG1B, PMI, PMI1; PMM1, PMM_1, PMMH-22, Sec53...; FPGT, GFPP; TSTA3, FX, P35B, SDR4E1, GMDS, GMD, SDR3E1, GMPPB, LGMDR19, MDDGA14, MDDGB14, MDDGC14...; PFKL, ATP-PFK, PFK-B, PFK-L...; TPI1, HEL-S-49, TIM, TPI, TPID; ALDOA, ALDA, GSD12, HEL-S-87p...; TKFC, DAK, NET45; FBP1, FBP...; FCSK, 1110046B12Rik, CDGF2, FUK; ENOSF1, FUCD, RTS, TYMSAS

OK

4.3 Integration of genes/proteins with MS peaks

4.3.1 Data Upload (genes and MS peaks)



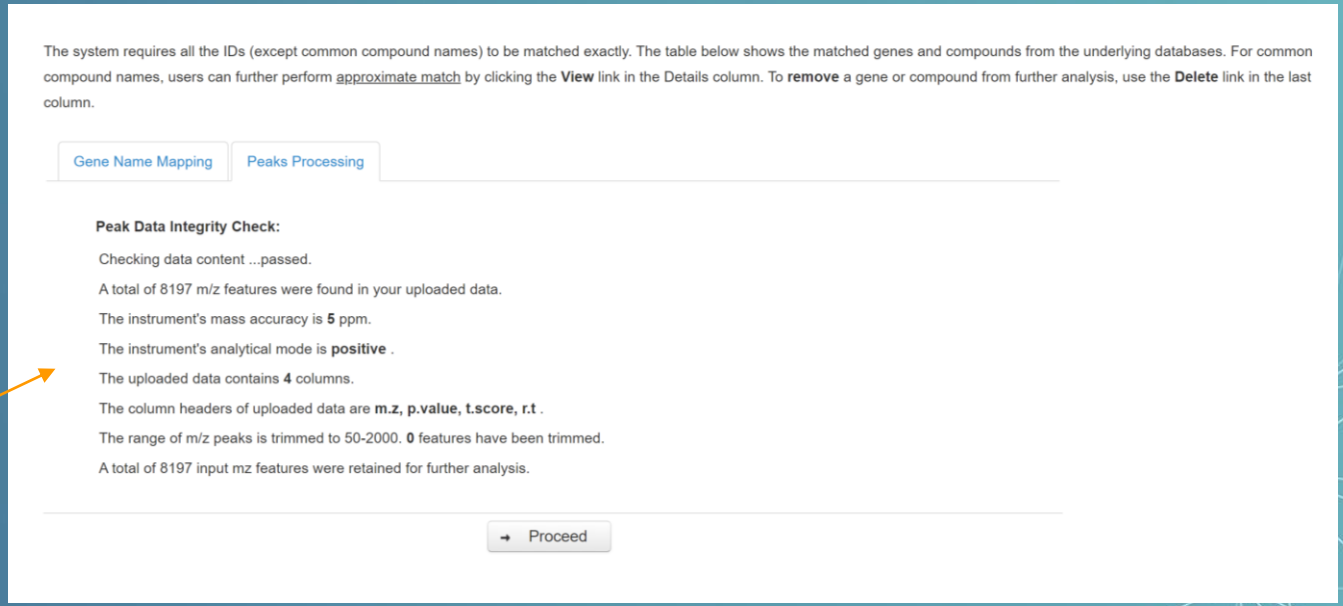
Change the metabolomics type as **'Untargeted (Peak list)'**

A peak list must be uploaded here

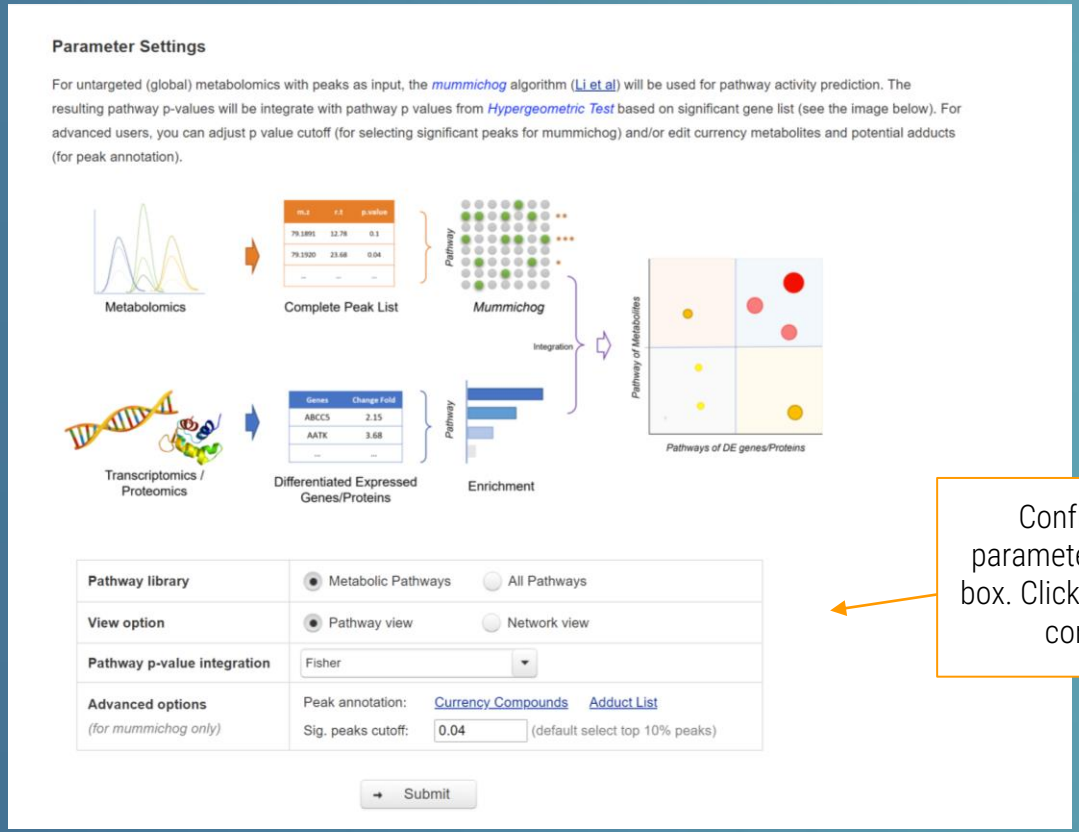
Click **Submit** to Proceed

4.3.2 Name Matching and peaks processing

Results of the name mapping of the uploaded gene to MetaboAnalyst's internal database, and peaks processing results are shown in this page. Click "**Proceed**" to continue.



4.3.3 Parameter Setting



TIP1: Pathway Library include two options. ‘Metabolic Pathways’ refers to the pathways include both genes and compounds, while ‘All Pathways’ also includes the pathways have containing genes only.

TIP2: P value cut-off is estimated automatically as the top 10% significant peaks.

Configure the parameters from this box. Click “**Proceed**” to continue.

4.3.4 Integration Results

Pathway Name	(M.,G.)/Total	p compound	p gene/protein	p merged	-log(p)	Holm p	FDR	Link
Citrate cycle (TCA cycle)	(2, 6)/46	0.858	6.2351E-8	1.0E-6	6.0226	1.0E-4	1.0E-4	KEGG
Tryptophan metabolism	(18, 5)/83	0.17932	1.2385E-5	3.1E-5	4.5068	0.0024	0.0012	KEGG
Fatty acid degradation	(2, 5)/80	0.99033	1.5635E-5	1.87E-4	3.7282	0.0144	0.0049	KEGG
Glyoxylate and dicarboxylate metabolism	(4, 4)/61	0.73371	6.2241E-5	5.02E-4	3.2992	0.0382	0.0099	KEGG
Lysine degradation	(5, 5)/80	0.75364	7.7877E-5	6.31E-4	3.2003	0.0473	0.01	KEGG

The result page is similar to 4.2.4. Different from the compounds uplaoding, the matched peaks will be shown in dark blue. The result label is showing the putative adduct for this peaks.

Thanks

If you have any questions please read through the FAQs or contact us at Zhiqiang.pang[at]xialab.ca or Jeff.xia[at]xialab.ca